AD_____

AWARD NUMBER: DAMD17-03-1-0136

TITLE: SRD5A1 Genetic Variation and Prostate Cancer Epidemiology

PRINCIPAL INVESTIGATOR: Troy Phipps

CONTRACTING ORGANIZATION: University of Southern California
Los Angeles, California 90033

REPORT DATE: May 2006

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-05-2006	Annual Summary	1 May 2003 – 30 Apr 2006

4. TITLE AND SUBTITLE

SRD5A1 Genetic Variation and Prostate Cancer Epidemiology

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-03-1-0136

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Troy Phipps

E-Mail: tphipps@usc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Southern California
Los Angeles, California 90033

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The human steroid 5-alpha reductase type I (SRD5A1) gene was sequenced in 101 men in order to identify genetic variants that might predispose carriers to prostate cancer. In these analyses, we uncovered a series of single nucleotide polymorphisms (SNPs) in the 3'-untranslated region (3'-UTR) of the SRD5A1 mRNA. However, we did not find SNPs that changed amino acid identities of the SRD5A1 protein. To further pursue the relevance of the SNPs in the 3'-UTR of this gene, we used cell culture assays to measure how they may alter RNA half-life, steady-state, and translation. Different combinations of 3'-UTR SNPs ("RNA haplotypes") resulted in reduced SRD5A1 mRNA half-life and steady state levels in our assay system. This data is consistent with the hypothesis that SNPs in the 3'-UTRs of mRNAs may play a role in the regulation of the SRD5A1 gene. Therefore, these SNPs should be considered as candidates for having biological function that might predispose to prostate cancer.

15. SUBJECT TERMS
Steroid 5-alpha reductase type I, prostate cancer, single nucleotide polymorphisms, SNPs, mRNA stability, haplotype

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	21	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

INTRODUCTION

In the 1940s, a link between androgens and prostate cancer was established based on the reduction or elimination of prostate carcinoma by castration as well as the observed opposite effect caused by injections of testosterone (a steroid hormone) (Huggins, C.B., Hodges, C.V. 1941). Later work found that steroid hormones bind to receptors and regulate nuclear gene regulation, thus altering cell and tissue phenotypes (Lefkowitz, R.J., et al. 1970; Attramadal, A., et al. 1976) Enzymes and biochemical pathways that control androgen metabolic flux are of significant interest to understanding prostate cancer predisposition studies because they participate in the regulation, distribution, and concentration of the various types of steroid hormones.

Prostate cancer is strongly influenced by genetics, with the estimated contribution being 42% of the risk (Lichtenstein, et al. 2000). Although most sporadic cancers are caused by environmental factors and *de novo* mutations, the relatively large genetic contribution in prostate cancer demonstrates how wide are the gaps in our knowledge and understanding of cancer genetics (Lictenstein, P., et al. 2000). Therefore, an examination of the different enzymes that catalyze the anabolism of testosterone into dihydrotestosterone (DHT) would be useful for investigating how increased metabolic flux through these pathways impact prostate growth and the likelihood of tumor development (Davies, et al. 1991; George, et al. 1991; Wilson, et al. 1975). There are two known steroid 5- alpha reductases in humans, types I and II. The two genes (*SRD5A1* and *SRD5A2,* respectively) have similar gene structure, and encode steroid 5-alpha reductases, but lie on different chromosomes and show different tissue expression. The type II reductase has been studied in relation to prostate cancer and certain SNPs play a role in prostate cancer predisposition for some populations.

The steroid 5-alpha reductase type II enzyme (E.C. 1.3.99.5), which catalyzes the reduction of testosterone (and other steroids) into the physiologically more active dihydrotestosterone (DHT) was shown to have at least ten single amino acid substitutions and three double mutations, all of which occur in normal, healthy males (Makridakis, et al. 2000). Biochemical analyses of these variant enzymes have a spectrum of activities that range from those having slower to those having faster reaction kinetics relative to the wild-type enzyme (Makridakis, N. et al. 1999; Makridakis, N., di Salle, E., Reichardt, J.K.V. 2000). One mutation, which changes an alanine at amino acid 49 into a threonine, boosts the chances of African-American men developing prostate cancer by 7.2-fold (Makridakis, et al. 1999). Meanwhile, the same mutation imparted Latino-American men with 3.6 times the likelihood of developing the disease (Makridakis, et al. 1999). Mutant enzymes also demonstrated a 32-fold range of inhibition to the selective steroid 5-alpha reductase inhibitor, finasteride (Makridakis, M, et al. 2004). In the prostate cancer prevention trial (PCPT), daily oral dosing of finasteride decreased prostate cancer risk by roughly 25%, while those that did develop cancer and who had taken finasteride were at a 6-fold higher risk for more aggressive, higher grade cancers (Thompson, IM, et al. 2003). This poses the question: Are there common sequence variants in the type I reductase that predispose men to or protect men from prostate cancer or benign prostatic hyperplasia (BPH)?

SPECIFIC AIMS

Specific Aim #1 involved uncovering sequence variation in the SRD5A1 gene in African-American, Latino-American, European-American, and Japanese-American populations.

Specific Aim #2 entailed identifying how SNPs alter steroid 5-alpha reductase type I function using biochemical and kinetic assays.

Specific Aim #3 involved a case-cohort study to determine if the SNPs uncovered increase or decrease the risk of developing prostate cancer in African-Americans, Latino-Americans, Japanese-Americans, and European-Americans, and if there is an interaction between the gene and environmental factors, such as dietary fat and obesity.

BODY

The *SRD5A1* gene encodes the steroid 5-alpha reductase enzyme type I which reduces testosterone into a physiologically more active chemical form, dihydrotestosterone (DHT). We hypothesize that the human steroid 5-alpha reductase type I enzyme plays a role in prostate cancer development, and that gain-of-function single nucleotide polymorphisms (SNPs) in the *SRD5A1* gene can increase androgen metabolic flux to the prostate, increasing its size, and thus predisposing to prostate cancer.

The overall goal in **Year One** was to identify and catalog the genetic variation in the *SRD5A1* gene. By direct sequencing analyses, SNPs were discovered throughout the transcribed portions (including 5'- and 3'-untranslated regions, UTRs) of the gene, the putative promoter region, and nearby splice junctions. **Table 1** lists 30 SNPS that were found in an initial genetic screen of genomic DNA isolated from 101 (out of 118 samples) ethnically diverse men. Nucleotide numbering of the SNPs is from the AUG codon based on data from NCBI accession number NM_000348. Two additional SNPs (1202T>C and 1578C>T) listed in the **Table 1** are not present in previous reports. Interestingly, it should be noted that no missense, frameshift, splice junction (3'-AG acceptor or 5'-GT donor sequences), or nonsense mutations were uncovered. Furthermore, all of the SNPs uncovered were in Hardy-Weinberg Equilibrium.

Thus far, there have been few (if any) systematic studies of the biological relevance groupings of SNPs (herein called "RNA haplotypes") in present in 3'-UTRs. Generally, one or two such SNPs are tested for an effect in an artificial expression vector (e.g. luciferase reporter constructs). Furthermore, to our knowledge, there are no published studies that focus on the effect of haplotypes on mRNA half-life, steady state levels, or the efficiency of translation. Here, we hypothesize that 'RNA haplotypes' constitute a class of genetic variation that plays an important role in regulating *SRD5A1* gene expression. This provides the rationale for the additional analysis on RNA biochemistry undertaken during the final year of the granting period. The specific aims in **Year Two** sought to answer fundamental questions as to how common RNA haplotypes present in a human population may affect the mRNA and protein levels of the *SRD5A1* gene.

Table 1. Single Nucleotide Polymorphisms Detected in the Human *SRD5A1* gene.

Gene Region	SNP	Genotypes	Observed Frequencies (Observed/Total)	Allele	Allele Frequency	Number of Chromosomes Sequenced
PROMOTER	-815G>A	G/G	46/85	G	0.72	170
		G/A	30/85	A	0.28	
		A/A	9/85			
	-781G>A	G/A	30/85	G	0.51	170
		G/G	28/85	A	0.49	
		A/A	27/85			
	-748T>G	T/T	31/85	T	0.55	170
		T/G	31/85	G	0.45	
		G/G	23/85			
	-648C>T	C/C	48/87	C	0.72	174
		C/T	30/87	T	0.28	
		T/T	9/87			
	-618A>G	A/A	82/86	G	0.02	172
		G/A	4/86	A	0.98	
		G/G	0/86			
	-571G>T	G/G	85/86	G	0.99	172
		G/T	1/86	T	0.01	
		T/T	0/86			
	-525G>A	G/G	82/85	G	0.98	170
		G/A	2/85	A	0.02	
		A/A	1/85			
	-477C>T	C/C	51/86	C	0.77	172
		C/T	31/86	T	0.23	
		T/T	4/86			
	-382delA	A/A	76/87	A	0.94	174
		A/del	11/87	del	0.06	
		del/del	0/87			
	-290A>G	A/A	86/87	A	0.99	174
		A/G	0/87	G	0.01	
		G/G	1/87			
	-278C>G	C/C	93/97	G	0.02	194
		G/C	4/97	C	0.98	
		G/G	0/97			
	-258G>A	G/G	82/87	G	0.97	174
		G/A	4/87	A	0.03	
		A/A	1/87			
	-243C>T	C/C	48/85	C	0.85	170
		C/T	24/85	T	0.15	
		T/T	13/85			
EXON 1	217G>C	G/G	31/101	G	0.51	202
		G/C	41/101	C	0.49	
		C/C	29/101			
INTRON 1	IVS1-17G>A	G/G	91/98	G	0.96	196
		G/A	7/98	A	0.04	
		A/A	0/98			

Table 1 (continued). SNPs Detected in the Human *SRD5A1* gene.

EXON 2	436A>G	G/G	4/99	G	0.27	198
		G/A	46/99	A	0.73	
		A/A	49/99			
	475G>A	G/G	56/99	G	0.76	198
		G/A	39/99	A	0.24	
		A/A	4/99			
INTRON 2	INV2+8T>C	T/T	95/98	T	0.98	196
		T/C	3/98	C	0.02	
		C/C	0/98			
	IVS2-30C>T	C/C	91/100	T	0.05	200
		T/C	8/100	C	0.95	
		T/T	1/100			
EXON 3	604A>G	A/A	98/100	G	0.01	200
		G/A	2/100	A	0.99	
		G/G	0/100			
	607G>A	G/G	57/100	G	0.76	200
		G/A	37/100	A	0.25	
		A/A	6/100			
EXON 4	706C>T	C/C	98/103	T	0.02	206
		T/C	5/103	C	0.98	
		T/T	0/103			
3'UTR	1202T>C	T/T	108/109	T	0.99	218
		T/G	1/109	G	0.01	
		G/G	0/109			
	1368C>T	T/C	40/100	T	0.43	200
		C/C	37/100	C	0.57	
		T/T	23/100			
	1441G>A	G/G	88/95	G	0.96	190
		G/A	6/95	A	0.04	
		A/A	1/95			
	1526A>G	A/A	88/99	G	0.07	198
		G/A	8/99	A	0.93	
		G/G	3/99			
	1535G>A	G/G	94/100	G	0.97	200
		G/A	6/100	A	0.03	
		A/A	0/100			
	1578C>T	C/C	92/99	C	0.93	198
		C/T	6/99	T	0.07	
		T/T	1/99			
	1758C>T	C/C	92/99	T	0.04	198
		T/C	6/99	C	0.96	
		T/T	1/99			
	2063T>C	T/T	99/103	T	0.98	206
		T/C	4/103	C	0.02	
		C/C	0/103			

Note: Color coding is provided for the sole purpose of highlighting genetic variation that occurs at each nucleotide position within the *SRD5A1* gene.

SPECIFIC AIM 1 RESULTS

- The open reading frame, UTRs, splice junction sequences, and 1-kb upstream of the *SRD5A1* gene were sequenced in a total of 101 genomic leukocyte DNA samples from African-American, Latino-American, European-American, and Japanese-American men. Half of the samples are from men who have been diagnosed with prostate cancer, and the other half from men without this disease. All samples were blinded to prevent ascertainment bias.

- Thirty different SNPs in the *SRD5A1* gene were discovered in the above population (see **Table 1** for the locations, genotypes, and allele frequencies of these SNPs)

- Eighty-eight percent of SNPs uncovered involved transitions (i.e. purine-purine or pyrimidine-pyrimidine substitutions). However, at nucleotide –382 of the promoter region, some men were heterozygous for a single base deletion (delA). We confirmed the presence of this deletion by restriction fragment length polymorphism (RFLP) analysis.

- The two SNPs (436A/G and 475G/A) present in exon 2 were found to be in strong linkage disequilibrium. SNPs at –618, -571, -477, -382, -290, -278, -258, 1202, 1578, and 1758 are not found in the NCBI database of single nucleotide polymorphisms (dbSNP found at http://www.ncbi.nlm.nih.gov/projects/SNP/).

SPECIFIC AIM 2 RESULTS

To begin accessing the function of SNPs in the 3'-UTR of the *SRD5A1* mRNA, we first investigated their phasing with respect to one another. Haplotypes were constructed from the genotyped cohort using the PHASE software package (hosted at http://archimedes.well.ox.ac.uk/). Haplotype blocks and linkage disequilibrium were calculated using haploBlockFinder software at the same site (**Table 2**).

Table 2. Sequences of Haplotype Blocks in *SRD5A1* 3'-UTR region.

	Haplotype Block #1					**Haplotype Block #2**				
	1202	1368	1441	1526	**Freq**	1535	1578	1758	2063	**Freq**
Alpha(wild-type)	T	C	G	A	**126**	G	C	T	T	**194**
Beta	T	T	G	A	**66**	G	T	T	T	**8**
Gamma	T	T	G	G	**13**	A	C	T	T	**6**
Omega	T	T	G	A	**9**	A	C	T	T	**6**
1535A	T	C	G	A	NA	A	C	T	T	NA
1368T	T	T	G	A	NA	G	C	T	T	NA
1578T	T	C	G	A	NA	G	T	T	T	NA

Note: Each haplotype block consists of four SNPs. Loci with the same sequence are grayed-out. Relative frequencies of each block are listed in the Freq column. Frequency indicates the number of chromosomes (N=214), while NA = not applicable.

Preliminary Computational Approaches

We first used a series of computational approaches to examine potential biological properties of SNPs and/or RNA haplotypes contained within the *SRD5A1* 3'-UTR. For example, we used TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) to predict differences in putative enhancer and transcription factor binding sites. The only major difference was that 1578U may create a c/EBP-1 transcription factor binding site. No predicted microRNA target sites were identified near any of the eight 3'-UTR SNPs. In contract, we uncovered potential differences in the secondary structures of the two full length *SRD5A1* mRNAs (**Figure 1**). Variation at the 1368 nucleotide site was predicted to result in major perturbations of RNA secondary structures. This is especially interesting given that RNA-binding proteins are known to specifically bind to the 3'-UTRs of mammalian mRNAs based on primary sequence as well as structural information (Ross, J. 1995; Di Paola, et al. 2002, Skalweit, A., et al. 2003).

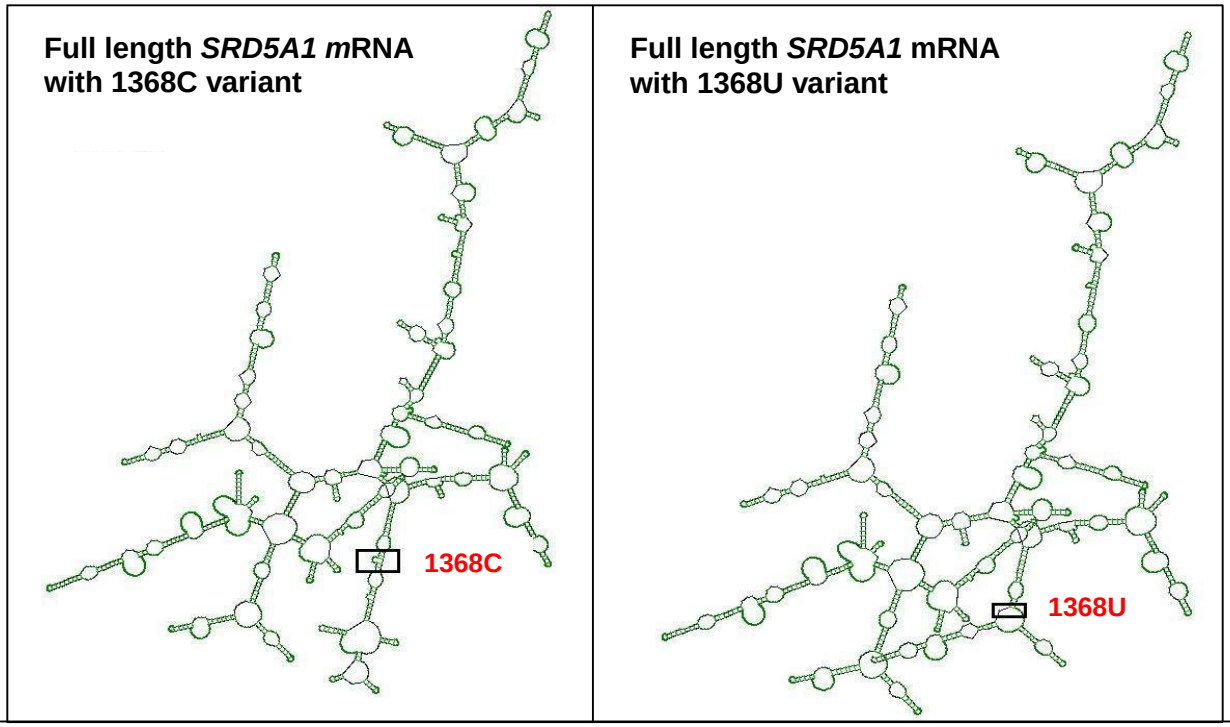

Figure 1. Predicted secondary structures of *SRD5A1* mRNAs. The Genebee software program (http://www.genebee.msu.su/) was used to predict the structures of *SRD5A1* mRNAs with either cytosine (left panel) or uracil (right panel) at nucleotide position 1368. A bifurcated stem loop structure may be formed proximal to 1368U (boxed region in left panel), while a bulged double helix is predicted to be formed proximal to 1368C (boxed region in right panel).

Functional Approaches to Elucidate the Relevance of SNPs in the *SRD5A1* 3'-UTR

Next, we created *in vitro* cell culture model systems in order to address the question if SNPs in the *SRD5A1* 3'-UTR can alter RNA stability and steady state levels in living cells. Two main classes of mammalian expression vectors were constructed. The first class of expression vectors contained luciferase reporter constructs fused to the full length *SRD5A1* 3'-UTR (**Figure 2**). The 3'-UTR was derived from a plasmid containing the entire *SRD5A1* cDNA and cloned downstream of the luciferase cDNA. The 3'-UTR sequences were then mutagenized to give rise to the sequence combinations listed in **Table 2**. All of the 3'-UTRs used in these experiments were engineered from the wild-type genetic background.

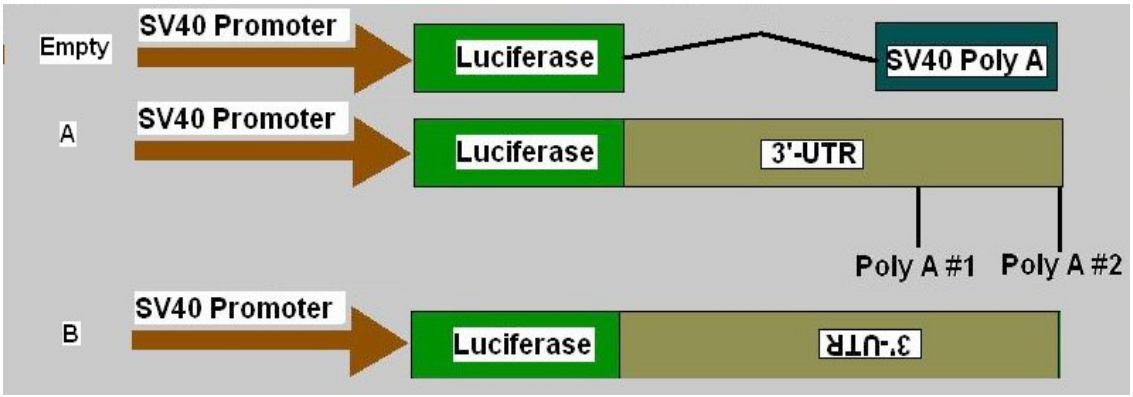

Figure 2. Schematic of the luciferase constructs used to transfect human embryonic kidney cells (HEK-293). Empty = empty pGL3P vector backbone. A = pGL3P+full length human *SRD5A1* 3'-UTR. B = pGL3P+inverted human *SRD5A1* 3'-UTR (denoted as UTR- in **Figure 4**). The 3'-UTR was excised from a plasmid containing the full length human *SRD5A1* cDNA, and mutagenesis was carried out on the luciferase constructs. Three independent clones from each mutagenesis reaction were used in transfections.

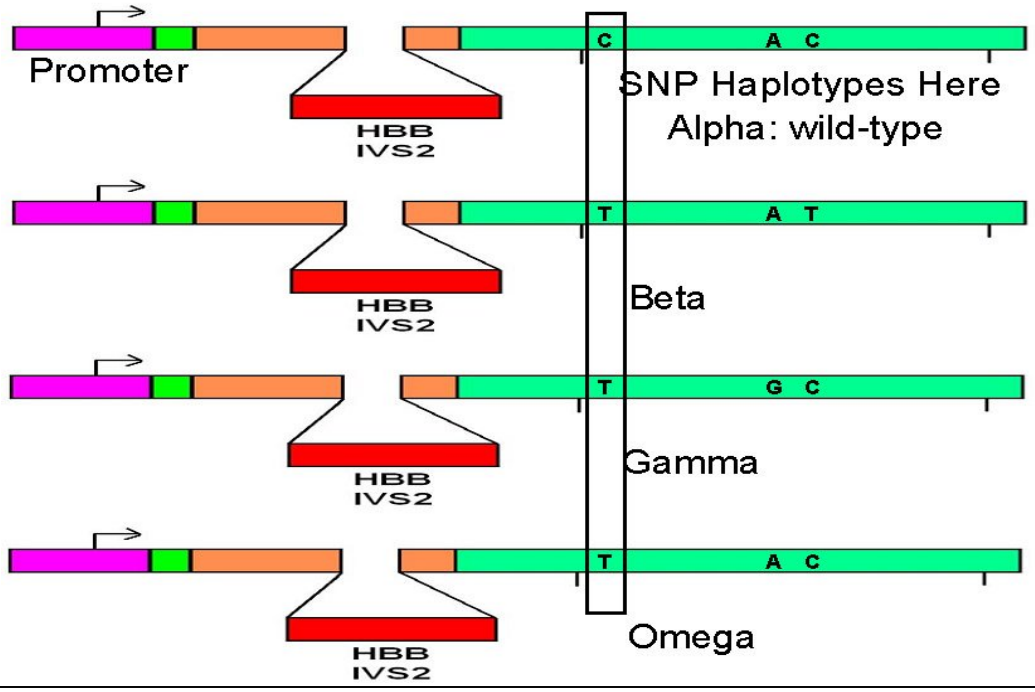

Figure 3. Diagram of *SRD5A1* expression constructs. An 850-bp region of the native *SRD5A1* promoter region was PCR amplified and cloned onto the full length 5'-UTR (green), cDNA (orange), with the addition of the human beta-globin intron 2 (PCR Amplified) near the site of the endogenous intron 4 (red), and 3'-UTR (green) containing both poly A signals. 3'-UTR SNPs were present as indicated by the boxed region. Only native endogenous restriction sites were used in vector construction and all viral sequences in the plasmid were removed.

The second class of vectors contained the native *SRD5A1* promoter, *SRD5A1* full length cDNA, full length 3'-UTR, and a human beta-globin intron near the site of the final intron in the native *SRD5A1* message (**Figure 3**). Introns are needed for proper export into the nucleus and poly-adenylation choice. When expressed, the intron was spliced out (sequenced after RT-PCR, data not shown), giving rise to an RNA message that contained the identical length and sequence as the chromosomally expressed *SRD5A1* mRNA. The promoter and introns were obtained via PCR, while the full length *SRD5A1* cDNA was from a plasmid kindly donated by Dr. David Russell (University of Texas, Southwestern). All transfections were transient and carried out in an African Green Monkey kidney cell line (COS-7). The COS-7 cells were harvested 48 hours post-transfection, sonicated, and total protein concentration was measured via a modified Bradford Assay. Luciferase signal was measured and beta-galactosidase enzyme assays were used to correct for transfection efficiency.

Beta, gamma, and omega haplotypes showed a 30-50% reduction of luciferase signal as compared to the wild-type (alpha) haplotype. **Figure 4** shows the luciferase data normalized to the wild-type, stars indicate statistically significant differences as determined by a two-tailed, equal variance Student's t-test. The "1368U" construct provided ~50% more signal than the wild-type "1368C" construct. Based on the beta, gamma, and omega sequences, there are likely to be 2-3 SNPs interacting with each other and leading to the observed data. The individual SNPs were evaluated on a wild-type background, but no single SNP alone was sufficient to show a 50% reduction in RNA levels.

It is puzzling how the 1368U construct yields approximately 4-fold higher luciferase signal relative to the beta, gamma, and omega haplotype constructs even though they all share the same uracil residue at nucleotide position 1368. One hypothesis is that some combinations of SNPs may alter RNA steady state levels, while others change RNA half-life, with the certain combinations showing increased steady-state RNA, higher luciferase, but a reduced half-life. Currently, we are investigating the effects of multiple SNP combinations, as well as SNPs at nucleotide positions 1535 and 1578.

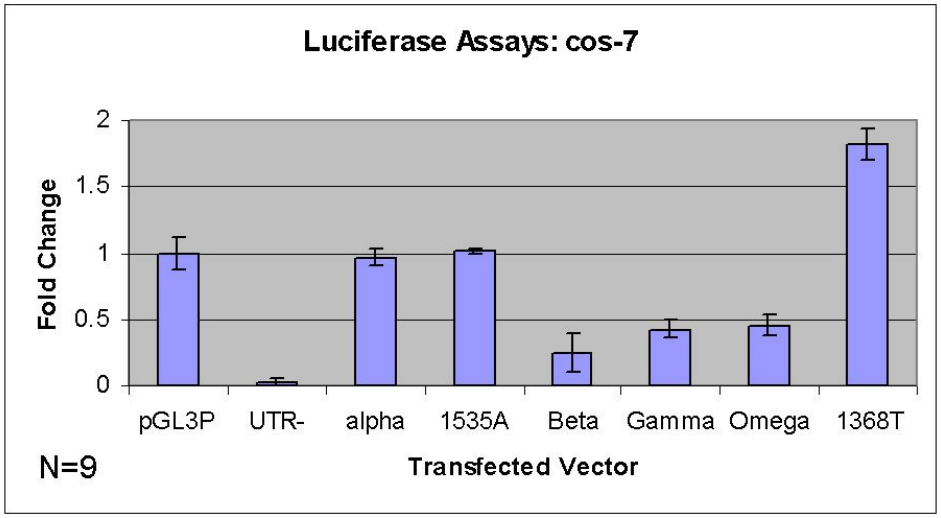

Figure 4. The luciferase clones containing different *SRD5A1* haplotypes were transfected into COS-7 cells. The data was corrected for total protein and transfection efficiency via β-galactosidase activity. All transfections were averaged after using triplicate clones replicated 3X.

To measure RNA steady state or half-life levels of the various haplotypes, relative reverse transcriptase polymerase chain reactions were used (RT-PCR). Endogenous *GAPDH* was used to normalize *SRD5A1* mRNA levels for comparison.

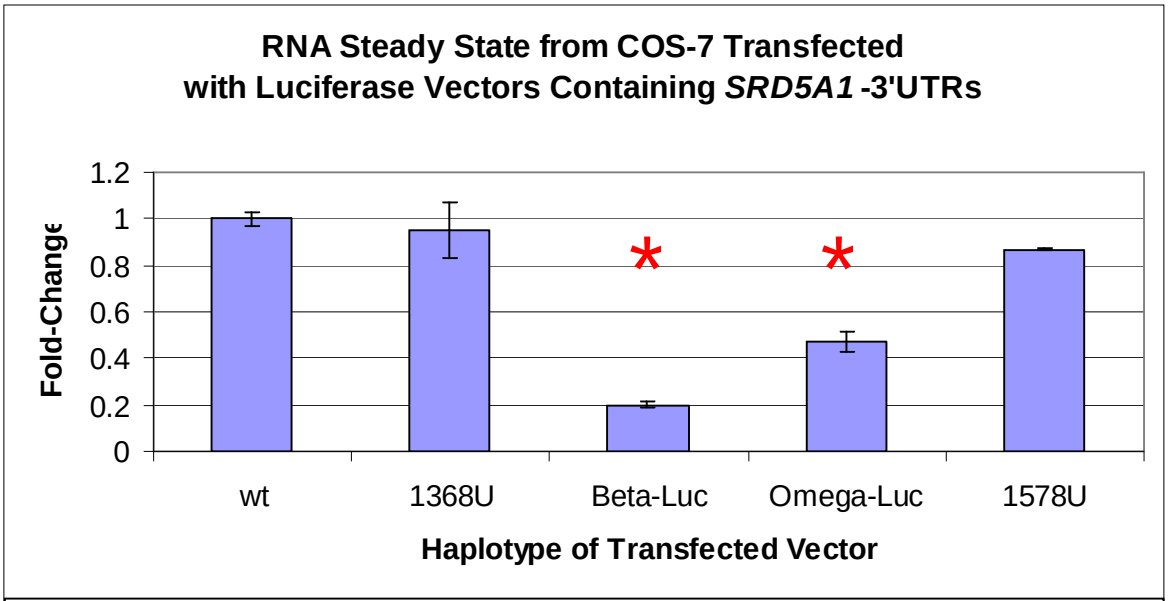

Figure 5. *SRD5A1* steady state RNA levels for different haplotypes after transfection and normalization using glyceraldehyde phosphate dehydrogenase (*GAPDH*). Stars represent statistically significant differences as indicated by using a two-tailed, equal variance t-test. N=9 (i.e. three clones transfected three times).

Figure 5 shows averaged data and standard errors for RNA steady state levels of several luciferase haplotypes transfected into COS-7 cells. Both beta and omega haplotypes had a reduced RNA steady state level compared to the wild-type, while 1368U and 1578U showed no difference. **Figures 6-7** and **Table 3** show RNA half-life experiments comparing both luciferase vectors next to native expression vectors. Therefore, SNPs at nucleotide positions 1368 and 1578 alone do not alter RNA half-life. Both luciferase and native expression vectors showed similar patterns of RNA half-life that matched the respective haplotypes. Thus, artificial constructs are suitable for assessing the properties of different 3'-UTR RNA haplotypes.

When a particular vector showed a reduced half-life, its RNA steady state was also reduced by a similar amount (e.g. beta). Moreover, reducing RNA levels also led to reduced luciferase activity. This suggests that the constructs with beta, omega, and gamma haplotypes show reduced activity due to reduced levels of RNA translation.

1368U demonstrated a similar RNA steady-state and half-life to the wild-type, but had increased luciferase signal. This suggests that the effect is not caused by an RNA structural change, but rather by a trans-factor, most likely a protein. To address this, HEK-293 cells were partially purified and used to identify the protein(s) that may be binding at this locus.

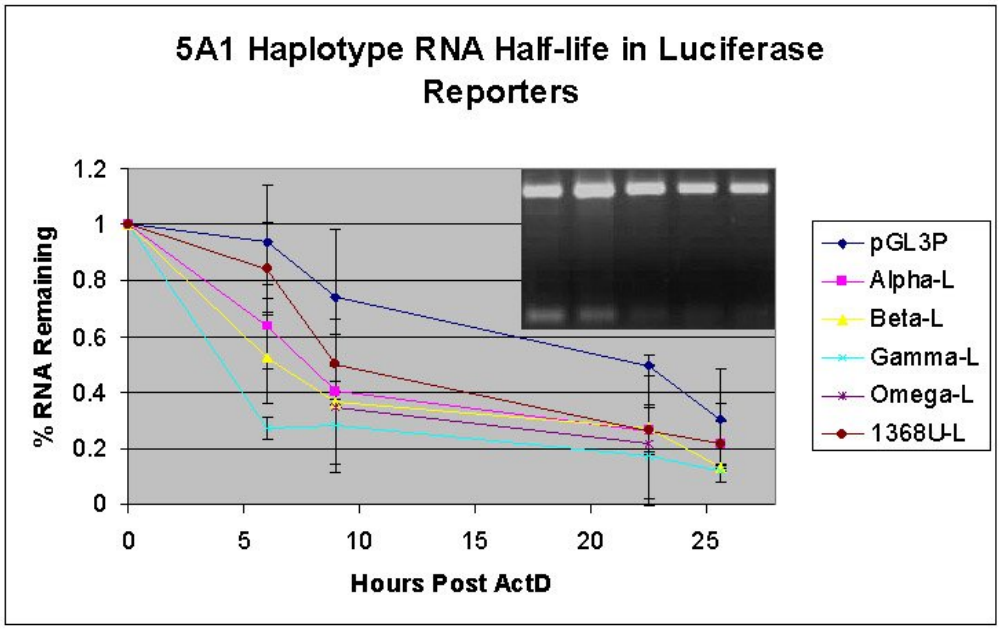

Figure 6. *SRD5A1* RNA half-life after actinomycin D treatment (transcriptional inhibitor). The decay of luciferase mRNA containing the indicated SRD5A1 3'-UTR haplotypes is shown. qRT-PCR was used to normalize *SRD5A1* expression against *GAPDH*. The larger band is *GAPDH*, and the smaller band is 5A1. The data is derived from three separate transfections of a single clone.

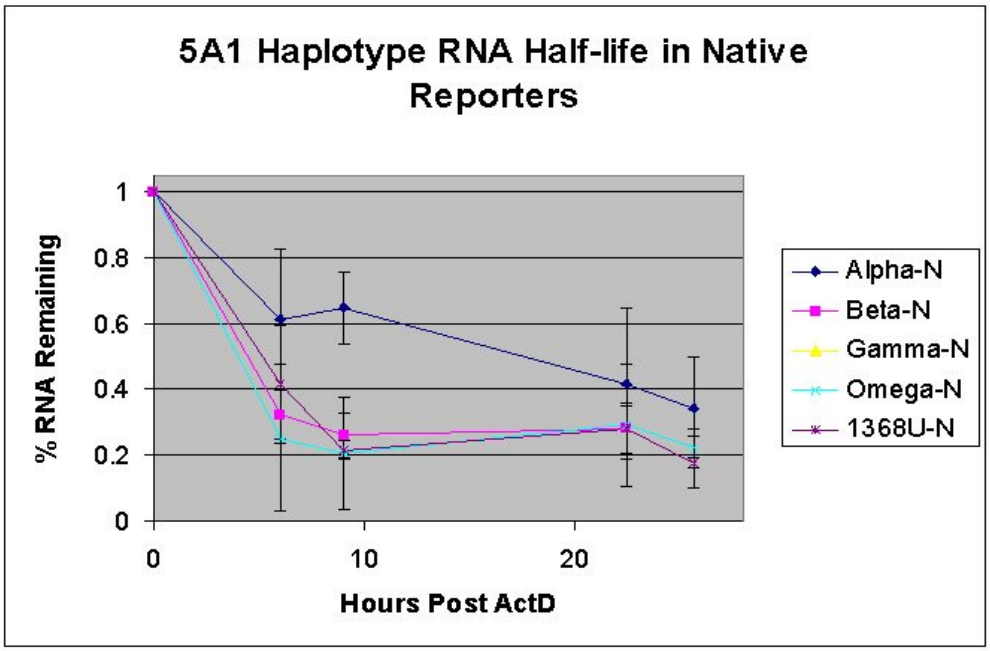

Figure 7. *SRD5A1* RNA half-life after actinomycin D treatment. The decay of *SRD5A1* mRNA produced from native *SRD5A1* expression plasmids is shown. As was the case in Figure 7, qRT-PCR was used to normalize *SRD5A1* expression against *GAPDH*. The data is derived from three separate transfections of a single clone.

Table 3. RNA half-lives for *SRD5A1* constructs transfected into 293 cells.

	Ave	Std Dev
pGL3P	15.3	3.8
Alpha-L	11.1	3.2
Alpha-N	11.1	3.0
Beta-L	5.4	1.1
Beta-N	3.7	0.7
Gamma-L	3.3	0.9
Gamma-N	4.2	1.1
Omega-L	5.9	NA
Omega-N	5.3	2.3
1368T-L	9.2	1.0
1368T-N	9.4	2.5

Note: L = luciferase containing plasmids, N = native SRD5A1 expression plasmids.

A 31-bp fragment of the *SRD5A1* 3'-UTR containing 1368C or T in the center (uracil will replace thymine when transcribed) was amplified with a normal reverse primer as well as a forward primer containing a SP6 promoter sequence appended to its 5'-end. Radioactive riboguanosine triphosphate was added to an *in vitro* transcription reaction and then bound to whole cell extracts, or fractions from a human kidney cell line (HEK-293). The bound RNA-proteins were separated on polyacrylamide gels, dried, and exposed to a phosphor screen. To estimate molecular weight, UV light was used to covalently cross-link the complex prior to running on a denaturing polyacrylamide gel.

To determine specific binding, self or non-self single stranded RNA 31-mers were added in molar excess. In **Figure 8**, both 1368C and 1368U oligos bind specific and non-specific proteins. Cross-competition with the opposite RNA showed that 1368U competed better for the specific protein than did 1368C. The 1368U RNA fully out-competed specific complexes at 1000x concentration, while the 1368C RNA only provided partial competition at the same concentration. This suggests that a protein complex may have improved affinity for the 1368U allele than to 1368C.

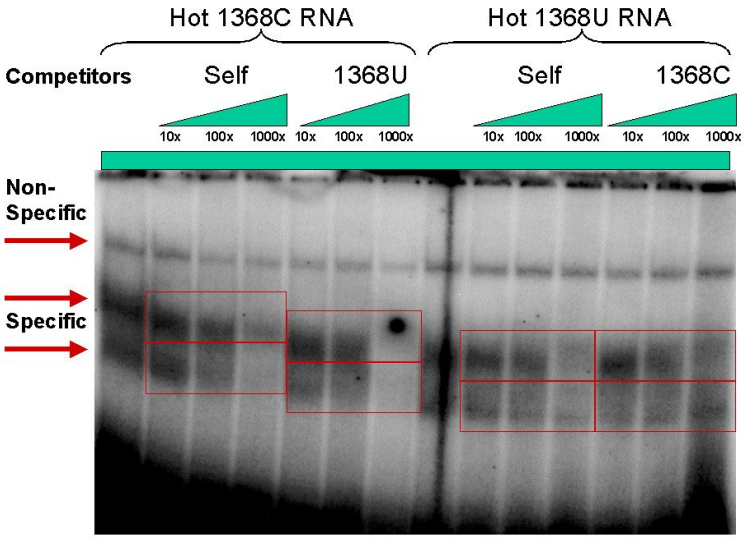

Figure 8.1368U competes better for specific protein binding than does 1368C. Native RNA Gel Shift using radiolabeled 1368C/U single stranded RNA and bound to whole cell HEK-293 extract was run on an acrylamide gel. Increasing amounts of self, or non-self competitors are indicated above the wells. Non-specific and specific bands are marked with red arrows, while specific competition bands are boxed in red.

An attempt to partially purify and identify the proteins via mass spectroscopy was undertaken using HEK-293 whole cell extracts. The first step was dialysis, followed by two ammonium sulfate precipitations (20% then 50%), with the final step being an RNA affinity column bound with 1368U single stranded RNA. Proteins of approximately 70kDa was isolated and sent to Stanford University's Mass Spectrometry Core Laboratory. Two proteins were identified: heat shock protein 70-1, and Ku70. Currently, experiments are underway to evaluate the binding of these proteins to different allelic variants of the *SRD5A1* 3'-UTR.

Specific Aim #3 Results

Unfortunately, the DNA samples used in the study were very difficult to PCR amplify and sequence due to their degraded state. Because of this technical difficulty, Specific Aim #3 was not completed as anticipated. During the third year, genotype troubleshooting along with *in vitro* biochemistry of the *SRD5A1* 3'UTR SNPs was conducted. The additional biochemical data was included in Specific Aim #2 for this final report, although the work was completed during the third year.

However, an interesting finding was revealed when the above genotyped samples were stratified at SNP 1368 based on ethnicity. This SNP was chosen for stratification because it lay in the *SRD5A1* 3'UTR, had the highest allele frequencies of both bases in the 3'UTR, and was barely in Hardy-Weinberg Equilibrium (HWE). The 101 LHMEC samples genotyped were composed of three different ethnicities: Japanese-Americans, Caucasians, and African-Americans. While both the Japanese-American and Caucasian individuals showed similar allele frequencies compared to each other, they demonstrated a striking difference compared to African-Americans (**Table 4**). A separate pure Japanese cohort of 616 men (YUSUKE in dbSNP) showed similar allele frequencies compared to the Asian-Americans and Caucasians in our genotyped LHMEC samples. This contrasted to the allele frequencies of the LHMEC African-American sub-population, which had an over-representation of the 1368T allele by a factor of two and was not in HWE. The allele frequency differences found in pure

Japanese, Japanese-American, and Caucasian men versus African-Americans was interesting, because the former have significantly lower incidence and mortality rates for prostate cancer than the latter. The ancestral human allele is 1368C as determined by comparing DNA sequences from Chimpanzees and Rhesus Monkeys which have perfect homology to the human *SRD5A1* 3'UTR around this SNP site. Since the 1368T allele arose only in human evolution and is overrepresented in African-Americans, this SNP was anticipated to be functional. Even if '1368T' is truly neutral, it may still be useful for prostate cancer admixture mapping (Freedman et al., 2006; Smith et al., 2004).

Since SNP 1368T was predicted to alter RNA structure, and in preliminary biochemical experiments demonstrated increased luciferase activity when cloned onto a luciferase reporter, with no increase with RNA levels, it may act through increasing protein translation. Since the 1368T allele is over represented in the high risk African-American men, it is possible that this SNP may play some role in prostate cancer etiology. However, further biochemical assays are currently being conducted to validate or refute the above hypothesis.

1368 Allele Frequency	**YUSUKE**	**LHMEC**	**LHMEC**	**LHMEC**	**LHMEC**	**North American PDR90**
Ethnicity	Japanese	Japanese-American	Caucasian	African-American	Combined	Combined
C	0.73	0.72	0.67	0.32	0.57	0.54
T	0.27	0.28	0.33	0.68	0.43	0.46
X^2	Unknown	1.60	2.30	4.10	3.40	<0.01
HWE	Unknown	Yes	Yes	No	Yes	Yes
2N=	1232	64	72	64	200	142

Table 4. Ethnic distribution of the 3'UTR SNP 1368 in Three Different Populations. The *SRD5A1* 3'UTR SNP, 1368, demonstrates similar allele frequencies for a pure Japanese population (YUSUKE from dbSNP), Japanese-Americans, and Caucasians from the LA/HI Multiethnic Cohort (LHMEC). LHMEC African-Americans demonstrated significant allele skewing compared to the other populations. The Japanese cohort comprised 752 anonymous, unrelated volunteers (of which 616 gave DNA), while the Los Angeles-Hawaiian Multiethnic Cohort sample is a mixed population consisting of 36 Caucasians, 32 African-Americans, and 32 Japanese-Americans (101 combined). A Mixed North American Cohort (PDR90 from dbSNP) from 71 men, contained an unknown mixture of the same ethnicities as the LA/HI Multiethnic Cohort. 2N is the number of genotyped chromosomes. X^2 is the chi-squared value, using a cut-off of 3.8 to evaluate if a population was in Hardy-Weinberg Equilibrium (HWE).

KEY RESEARCH ACCOMPLISHMENTS

1) Delineation of *SRD5A1* SNPs and haplotypes present in a human cohort.

2) Completed functional testing of *SRD5A1* 3'-UTR SNPs in a luciferase reporter, and showed that the human *SRD5A1* 3'-UTR stabilizes luciferase signal while a random piece of DNA (UTR^-) ablates this effect to background.

3) Both luciferase and native vectors demonstrated similar RNA half-lives, and RNA steady states for the haplotypes, indicating that chimeric vectors are suitable to use in experiments.

4) The 1368C/U polymorphism is anticipated to alter RNA secondary structure. The combined data suggests that the increase in luciferase signal for the 1368U transcript is due to a trans-factor leading to increased protein tranlsation. Both specific and non-specific proteins bind to this region, although the protein(s) are still unkown.

5) The 1368T SNP within the *SRD5A1* 3'UTR was found in HWE when all ethnicities were combined; however, when analyzed by ethnicity, African-American men had approximately a 2 fold over representation of the 1368T SNP as compared to pure Japanese, Japanese-Americans, or Caucasians. The 1368T SNP may be useful for admixture mapping of diseases such as prostate cancer.

6) The beta, omega, and gamma haplotypes showed reduced RNA steady-state, RNA half-lives, and luciferase light, consistent with increased RNA decay, thus demonstrating that dramatic effects at the RNA and protein level can occur by mutating a small number of RNA bases in a message.

This work has bearing on prostate cancer because the SNPs tested occur naturally in men. The data suggests that SNPs that do not change amino acids or truncate proteins may have a real effect in terms of protein production, and thus on hormone catabolism.

REPORTABLE OUTCOMES

Troy Phipps will receive a Ph.D. degree and obtain a post-doctoral position based on data/experience generated from this award. Dr. Juergen Reichardt served as primary research mentor for over 95% of the granting period. As a colleague of Dr. Reichardt at USC, Dr. Hacia has agreed to oversee Mr. Phipps' research for the duration of his graduate career. The data shown in this Progress Report is being drafted into manuscript form, and will be submitted to an academic journal for peer review and publication in 2007.

CONCLUSIONS

Thirty SNPs in the human steroid 5-alpha reductases type I (*SRD5A1*) gene were discovered in a screen of 101 ethnically diverse men. One third of the SNPs are not found in the comprehensive SNP database (dbSNP). Although a number of SNPs were found scattered throughout the gene, none resulted in differences of the amino acid sequence of the *SRD5A1* protein. Therefore, we focused on genetic variation that occurred in the 3'-UTR of the *SRD5A1* transcript. This represents an original research direction since, to date, little is known about the effect SNPs and RNA haplotypes in the 3'-UTR of human mRNAs have on the transcript stability, half-life, and translation. On-going efforts are attempting to determine if molecular 'rules' concerning the biological properties RNA haplotypes exist. If so, these lessons can be extended to other messages and are likely to help substantively add to our collective knowledge of prostate cancer etiology via RNA biochemistry and metabolism.

REFERENCES

Attramadal, A., *et al*. Androgen Receptors in Male Sex Tissues of Rats and Humans. *Prog Clin Biol Res* **6**:189-203, 1976.

Chen, J.-M., Ferec, C., Cooper, D.N. A Systematic Analysis of Disease-Associated Variants in the 3' Regulatory Regions of Human Protein-Coding Genes I: General Principles and Overview. *Human Genet,* Epub ahead of print, 2006.

Davies, P., *et al*. Regulation of Prostate Growth. *J Endrocrinol* **131**: 5-17, 1991.

Di Paola, R., *et al*. A Variation in 3' UTR of hPTP1B Increases Specific Gene Expression and Associates with Insulin Resistance. *Am J Hum Genet* **70**: 806-812, 2002.

Freedman, M.L., Haiman, C.A., Patterson, N., McDonald, G.J., Tandon, A., Waliszewska, A., Penney, K., Steen, R.G., Ardlie, K., John, E.M.*, et al.* (2006). From the Cover: Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proceedings of the National Academy of Sciences of the United States of America *103*, 14068-14073.

George, F., Russell, D., and Wilson, J. Feed-forward Control of Prostate Growth: Dihydrotestosterone Induces Expression of Its Own Biosynthetic Enzyme, Steroid 5-alpha Reductase. *PNAS, USA* **88**: 8044-8047, 1991.

Huggins, C.B., Hodges, C.V. Studies on Prostate Cancer. I. The Effect of Castration, of Estrogen and of Androgen Injection on Serum Phosphatases in Metastatic Carcinoma of the Prostate. *Cancer Research* **1**:293-297, 1941.

Lefkowitz, R.K., *et al*. ACTH Receptors in the Adrenal: Specific Binding of ACTH-^{125}I and Its Relation to Adenyl Cyclase. *PNAS, USA* **65**: 745-752, 1970.

Lichtenstein, P., Holms, N.V., *et al*. Environmental and Hertiable Factors in the Causation of Cancer. *NEJM* **343**: 78-85, 2000.

Makridakis, N., di Salle, E., Reichardt, J.K.V. Biochemical and Pharmacogenetic Dissection of Human Steroid 5-alpha Reductase Type II. *Pharmacogenetics* **10**: 407-413, 2000.

Makridakis, N., *et al*. Association of Mis-sense Substitution in SRD5A2 Gene with Prostate Cancer in African-American and Hispanic Men in Los Angeles, USA. *Lancet* **354**: 975-978,1999.

Makridakis, N., Akalu, A., Reichardt, R.K.V. Identification and Characterization of Somatic Steroid 5-alpha Reductase (SRD5A2) Mutations in Human Prostate Cancer Tissue. *Oncogene* **23**: 7399–7405, 2004.

Ross, J. mRNA Stability in Mammalian Cells. *Microbiological Reviews* **59**:423-450, 1995.

Skalweit, A., *et al*. Posttranscriptional Control of Renin Synthesis: Identification of Proteins Interacting with Renin mRNA 3'UTR. *Circ Res* **92**: 419-427, 2003.

Smith, M.W., Patterson, N., Lautenberger, J.A., Truelove, A.L., McDonald, G.J., Waliszewska, A., Kessing, B.D., Malasky, M.J., Scafe, C., Le, E.*, et al.* (2004). A high-density admixture map for disease gene discovery in african americans. American journal of human genetics *74*, 1001-1013.

Thompson, I.M., Goodman, P.J., Tangen, C.M., Lucia, M.S., Miller, G.J., Ford, L.C., *et al*. The Influence of Finasteride on the Development of Prostate Cancer. *NEJM* **349**: 211–220, 2003.

Thompson, I.M., Klein, E.A., Lippman, S.M., Coltman, C.A., Djavan, B. Prevention of Prostate Cancer with Finasteride:US/European Perspective European. *Urology* **44**: 650–655, 2003.

Wilson, J. *Handbook of Physiology*, eds. Astwood, E. B. & Greep, R.O. (Am Physiol Soc, Washington), Vol **5**, Sect 7, Chap 25:491-508, 1975.